Prediction of Henry's Law Constant of Benzene Derivatives Using Quantum Chemical Continuum-Solvation Models

GERRIT SCHÜÜRMANN

Department of Chemical Ecotoxicology, UFZ Centre for Environmental Research, Permoserstr. 15, 04318 Leipzig, Germany

Received 28 February 1998; accepted 24 August 1999

ABSTRACT: Semiempirical (SM2, SM5.4A, MST-AM1, COSMO-AM1) and *ab inito* (HF/PCM-vdW, MP2//PCM-vdW, COSMO-DFT) dielectric continuum-solvation models as well as the surface-tension model SM5.0R are analyzed with respect to predicting Henry's law constant at 25°C using a compound set of benzene and 39 benzene derivatives. Both hydrophilic and hydrophobic compounds are covered with a total variation in Henry's law constant of almost eight orders of magnitude corresponding to 44 kJ/mol, and the data set is selected such that there are cases where subtle changes in the molecular structure result in substantial changes of the free energy of solvation. The calculations with SM2, COSMO-AM1, and COSMO-DFT include solution-phase geometry optimization, and the *ab initio* results refer to polarized basis sets of double-zeta quality, with two gradient-corrected functionals (BPW and BLYP) being used for the DFT-based models. The results show considerable differences in performance between the different continuum-solvation models, and among the methods yielding solvation free energies the systematic error ranges from −0.9 kJ/mol (SM5.0R) to 12.1 kJ/mol (MP2//PCM-vdW). In particular, the nonelectrostatic solvation energy contributions of SM2, SM5.4A, MST-AM1, and PCM-vdW do not correlate with each other, and with PCM-vdW omission of the nonelectrostatic component significantly improves the relative trend. The best statistics after scaling through linear regression are achieved with the electrostatic component of MP2//PCM-vdW ($r^2_{adj} = 0.94$) and with COSMO-DFT ($r^2_{adj} = 0.93$). The discussion includes detailed analyses of pecularities associated with certain functional groups, deviations from the expected relationship between dipole moment and solvation energy, and a simple approach to model dispersion interaction and cavitation energy by surface area

This article is dedicated to Prof. Dr. Martin Klessinger, Münster, Germany, on the occasion of his 65th birthday.

Correspondence to: G. Schüürmann; e-mail: gs@uoe.ufz.de

Journal of Computational Chemistry, Vol. 21, No. 1, 17–34 (2000)

CCC 0192-8651 / 00 / 010017-18

terms that differentiate between individual atom types. J Comput Chem 21: 17–34, 2000

Keywords: solvation; Henry's law constant; COSMO; PCM; SMx; MST-AM1

Introduction

The development of quantum chemical continuum-solvation methods has opened new ways to calculate chemical equilibria and reactions in aqueous solution.[1, 2] Recent examples include the computational analysis of S_N2 reaction profiles and *cis-trans* isomerization processes,[3] catalytic and bulk effects on the proton transfer with formamide,[4] and the prediction of the pK_a of organic acids.[5–7] The latter is also of interest for the evaluation of the biological activity of chemicals, which may directly depend on the acidity of the compound as was demonstrated for the uncoupling of the oxidative phosphorylation.[8]

For the area of environmental chemistry, an important compound property is the partition coefficient between air and water at thermodynamical equilibrium known as Henry's law constant,[9]

$$H = RT \cdot \frac{c_g}{c_w} \approx \frac{P_v}{c_w} \tag{1}$$

which is directly related to the free energy of aqueous solvation, ΔG_s, according to

$$\Delta G_s = RT \cdot \ln \frac{c_w}{c_g} = 2.3RT \cdot \log \frac{H}{RT} \tag{2}$$

where c_g and c_w denote the compound concentration in the gas phase and water phase, respectively, and P_v is the partial pressure of the compound in the gas phase. Recent reports of experimental H values for different groups of compounds as well as current attempts to improve or newly develop measurement techniques[10, 11] reflect the considerable interest in this thermodynamic compound property.

Only few increment methods are available to predict $\log H$ from chemical structure as reviewed recently,[12] and none of these methods has reached the precision and range of applicability known for corresponding methods to calculate the octanol/water partition coefficient as another environmentally relevant compound property. A more fundamental approach to calculate $\log H$ is given by its relationship to ΔG_s [eq. (2)], which in turn is available from molecular structure calculations that include a formal treatment of the effect of aqueous solvation.

For a small set of eight benzene derivatives covering the substituents CH_3, OH, OCH_3 and CN, the free energy perturbation method as a classical microscopic solvation model and the semiempirical continuum-solvation method SM2[13] yielded average errors in ΔG_s of about 2 and 3 kJ/mol, corresponding to factors of 2.2 and 3.6 for Henry's law constant.[14] More recently, the UAHF-PCM parameterization[15] of dielectric continuum-solvation models (DCMs) gave quite impressive results for 40 mainly small and polar solutes and 28 ionic solutes with mean errors in ΔG_s of around 0.8 and 4.2 kJ/mol.

However, there is very little experience about the performance of DCMs to predict ΔG_s and the associated Henry's law constant for more hydrophobic compounds like higher chlorinated benzenes, biphenyls and pesticides as well as other more persistent xenobiotics. With such compound classes, the nonelectrostatic contribution to ΔG_s is expected to play a greater role, which in turn may help refining current models to account for the cavitation energy and the dispersion-repulsion interaction.

In a previous investigation,[16] a corresponding analysis of Henry's law constants for benzene and 16 benzene derivatives was performed using SM2. Although both coefficient (0.169) and intercept (6.45) of the resultant regression relationship were close to the theoretical values according to

$$\log H = 0.175\Delta G_s + 6.39 \tag{3}$$

with ΔG_s in kJ/mol (referring to the standard state of 1 mol/L in the gas phase and in solution), and H in Pa*L/mol at 25°C, the overall prediction performance was only moderate, as indicated by a squared correlation coefficient (adjusted for degrees of freedom), r^2_{adj}, of 0.83, and a standard error (SE) of 0.65 log units. For these 17 compounds, MST-AM1[17] (using the Pauling set of atomic radii except for polar hydrogen with 0.9 Å, and with a scaling factor of 1.20 as recommended for semiempirical calculations) and COSMO-AM1[18] still show inferior results with r^2_{adj} values of 0.68 and 0.44, respectively.

It is tempting to explain this only limited success of SM2, MST-AM1, and COSMO-AM1 for benzene derivatives by the semiempirical level of calculation. To address this question more thoroughly, both

semiempirical and *ab initio* DCMs have been comparatively analyzed for their performance to predict ΔG_s and Henry's law constant of an extended set of 40 aromatic compounds including both hydrophobic congeners like xylene, dichlorobenzene and dibromobenzene as well as compounds with greater water solubilities like phenol, aniline and hydroquinone. A further selection criterion for the test set was the restriction to one aromatic ring as the parent structure to avoid additional problems of DCMs associated with ring structures[19] that may be caused by an improper handling of the entropic contribution to the free energy of solvation; entropy favors extended structures compared to cyclic systems.[20]

In addition to SM2, MST-AM1, COSMO-AM1, and the more recent SM5.4A[21] and SM5.0R[22] models, COSMO has been applied in the DFT (density functional theory) implementation[23] with atomic radii according to the recent COSMO-RS parameterization,[24] and an *ab initio* implementation of the PCM (polarizable continuum model) family using standard van der Waals radii[25] (PCM-vdW) was employed. The results show considerable variations in the prediction performance, which holds also true for the different *ab initio* DCMs. Moreover, the current schemes to address nonelectrostatic components to the solvation energy do not correlate with each other, indicating a need for improvement, which is particularly important for hydrophobic chemicals.

Materials and Methods

Decadic logarithms of experimental Henry's law constant (H in Pa*L/mol at 25°C) and associated solvation-free energies (ΔG_s in kJ/mol) of benzene and 39 benzene derivatives have been collected from various literature sources[16, 26–38] and are listed in Table I. For 22 compounds, H was calculated from experimental vapor pressure (P_v) and water solubility (S_w) at 25°C as P_v/S_w, and in a few cases P_v at 25°C has been interpolated from experimental values at higher and lower temperatures T (mostly 30 and 25°C) from plots of $\log P_v$ against $1/T$.

Initial three-dimensional geometries of the chemical structures were generated using the SYBYL molecular modelling package.[39] Subsequent quantum chemical calculations in the gas phase and in (simulated) aqueous solution were performed using MOPAC 93[40] (AM1, MST-AM1,[17] COSMO-AM1[18]), AMSOL[41] (SM2,[13] SM5.4A[21]), Gaussian 94[42] (HF/6-31G**, MP2//6-31G**, PCM-vdW[25, 43]), and DMol[44] (DFT, COSMO-DFT[23]). The molecular geometries were optimized at the respective levels of theory with the following exceptions: MST-AM1 energies refer to AM1 geometries, and PCM-vdW//HF as well as MP2 and PCM-vdW//MP2 energies refer to gas-phase geometries optimized at the HF/6-31G** level. For the surface-tension model SM5.0R[22] as implemented in AMSOL,[41] AM1-optimized geometries were used.

All energy values of PCM-vdW were calculated using the normalization procedure for the escaped charge with additional surface charges distributed according to the solute electronic density (option ICOMP = 4 of the PCM module).[45] However, this method was not available for the calculation of MP2-level solution-phase dipole moments, where the simpler normalization procedure using a constant factor (ICOMP = 2)[45] was applied.

For all DFT calculations, a numerical basis set with double-zeta quality including polarization functions on all atoms (dnp) was used together with the highest available option for the resolution of the integration grid (option xfine of DMol). Becke's gradient-corrected exchange functional[46] was combined with either the correlation functional of Perdew and Wang[47] or the one of Lee, Yang, and Parr[48] to yield BPW and BLYP, respectively. For recent discussions of the quality of these and other DFT models, the reader is referred to the literature.[49, 50]

A key parameter of continuum-solvation models are the atomic radii that are needed to define the molecular surface at which solute-solvent interaction takes place. The values for SM2, SM5.4A, and COSMO-AM1 are built in the software packages, and for MST-AM1 and PCM-vdW the (slightly modified) Pauling set (H (nonpolar): 1.20 Å, H (polar): 0.9 Å, C: 1.5 Å, N: 1.5 Å, O: 1.4 Å, Cl: 1.8 Å, Br: 2.0 Å, I: 2.15 Å, S: 1.80 Å) was combined with scaling factors of 1.20 (semiempirical) and 1.25 (*ab initio*) according to literature recommendations.[17, 25] The COSMO-DFT calculations included an outlying-charge correction,[51] and were performed with radii published recently for the COSMO-RS model, while the missing radii of Br, I and S were estimated as 115% of the respective Bondi-vdW radii (H: 1.30 Å, C: 2.00 Å, N: 1.83 Å, O: 1.72 Å, Cl: 2.05 Å, Br: 2.13 Å, I: 2.28 Å, S: 2.07 Å).

Finally, molecular surface areas and their contributions from individual atom types were calculated for the COSMO-DFT/BPW-dnp geometries, using MOLSV[52] and the above-mentioned radii as employed for that continuum-solvation model.

TABLE I.

Compound List with Henry's Law Constant and Free Energy of Solvation.[a]

		Experimental		DCM-AM1 and Surface Tension Model				PCM-vdW//6-31G**	
No.	Compound	log H	ΔG_S	ΔG_S^{SM2}	$\Delta G_S^{SM5.4A}$	$\Delta G_S^{SM5.0R}$	ΔG_S^{MST}	ΔG_S^{SCF}	ΔG_S^{MP2}
1	Benzene	5.75	−3.7	−2.3	−4.1	−3.4	−5.0	2.0	4.9
2	Toluene	5.83	−3.2	−1.2	−3.7	−1.7	−5.8	8.2	11.0
3	Phenol	1.60	−27.4	−24.2	−27.4	−26.5	−13.8	−20.7	−16.5
4	Chlorobenzene	5.57	−4.7	−4.8	−4.9	−3.6	−1.4	−0.8	2.4
5	1,2-Dichlorobenzene	5.28	−6.3	−7.1	−6.9	−4.0	0.6	−3.9	0.1
6	1,4-Dichlorobenzene	5.39	−5.7	−5.8	−4.2	−4.0	2.9	−1.2	1.9
7	Bromobenzene	5.40	−5.7	−8.3	−6.9	−4.2	−1.6	−0.2	3.2
8	1,2-Dibromobenzene	5.02[b]	−7.8	−14.3	−10.8	−5.0	−0.1	−0.7	2.6
9	1,4-Dibromobenzene	5.37[b]	−5.8	−13.4	−8.0	−5.0	3.4	0.9	4.0
10	Aniline	2.22[b]	−23.8	−28.3	−19.8	−19.8	−15.5	−19.2	−15.8
11	Anisol	4.16	−12.7	−9.8	−12.8	−13.9	−10.3	2.9	6.7
12	Benzaldehyde	3.43	−16.9	−19.7	−18.4	−23.0	−14.7	−12.8	−2.3
13	*o*-Cresol	2.08	−24.6	−20.5	−22.8	−24.0	−12.9	−12.0	−8.2
14	*m*-Cresol	1.93[b]	−25.5	−23.3	−27.2	−24.9	−14.4	−13.7	−9.5
15	*p*-Cresol	1.90	−25.7	−23.0	−26.7	−24.9	−14.3	−13.9	−10.0
16	Nitrobenzene	3.39	−17.1	−21.9	−12.8	−19.6	−17.2	−19.7	−8.4
17	Thiophenol	4.52	−10.7	−13.4	−13.0	−12.2	−8.4	−10.3	−7.9
18	*p*-Xylene	5.88	−2.9	0.0	−3.3	−0.1	−9.2	12.3	15.8
19	Hydroquinone	−1.50[b]	−45.1	−45.8	−50.5	−50.0	−28.7	−44.3	−38.6
20	Benzonitrile	3.41[b]	−17.0	−18.2	−12.4	−22.7	−12.4	−15.6	−8.2
21	1,2,3-Trimethoxybenzene	2.44[b]	−22.6	−17.0	−18.4	−23.2	−22.0	5.2	11.0
22	Fluorobenzene	5.80[b]	−3.4	−0.3	−0.9	−1.4	−3.7	−2.7	1.4
23	2-Chlorotoluene	5.55	−4.8	−3.9	−4.3	−2.2	−3.4	5.9	9.2
24	3-Chlorotoluene	6.21[b]	−1.0	−3.8	−4.3	−2.1	−3.7	5.7	9.0
25	Iodobenzene	5.10[b]	−7.4	−14.4	−6.9	−10.3	−2.8	−4.6	−0.8
26	1,2-Benzenediol	0.80[b]	−31.9	−40.2	−38.8	−42.5	−22.0	−35.9	−31.1
27	1,3-Benzenediol	−0.70[b]	−40.5	−45.4	−52.6	−49.5	−29.4	−45.4	−39.5
28	*o*-Phenylenediamine	1.12[b]	−30.1	−39.0	−30.8	−36.8	−34.5	−35.0	−30.6
29	*m*-Phenylenediamine	−1.13[b]	−43.0	−54.1	−35.3	−36.4	−34.1	−43.4	−38.6
30	*o*-Toluenenitrile	3.12[b]	−18.7	−16.3	−11.8	−21.6	−12.8	−7.4	−0.2
31	2,6-Dichlorobenzonitrile	2.86[b]	−20.2	−20.4	−16.8	−24.0	−4.2	−15.8	−8.5
32	2-Nitrotoluene	3.73[b]	−15.2	−17.4	−12.1	−17.6	−22.1	−10.9	−0.8
33	2,4-Dinitrotoluene	1.00[b]	−30.8	−30.7	−13.4	−33.7	−30.2	−28.8	−10.3
34	2,4,6-Trinitrotoluene	0.27[b]	−35.0	−35.7	−2.9	−49.4	−33.5	−40.3	−13.8
35	*m*-Chloronitrobenzene	3.97[b]	−13.8	−21.4	−10.3	−19.9	−15.6	−17.2	−6.6
36	*m*-Bromonitrobenzene	2.27[b]	−23.5	−24.7	−12.0	−20.4	−15.3	−16.0	−5.5
37	Acetophenone	3.03	−19.2	−18.5	−19.5	−22.3	−21.5	−8.3	2.0
38	2-Methylaniline	2.39[b]	−22.9	−20.9	−18.5	−18.6	−22.6	−14.0	−9.9
39	*N*-Methylaniline	3.06[b]	−19.0	−18.8	−17.9	−14.8	−19.9	−8.3	−4.1
40	*N*,*N*-Dimethylaniline	3.85[b]	−14.5	−12.7	−13.6	−8.2	−16.7	11.4	15.4
Average		3.3	−17.7	−19.0	−15.9	−18.7	−13.7	−11.7	−5.6
Range of variation		7.7	44.1	54.1	51.7	49.9	37.9	57.7	55.3
Average error		—	—	−1.3	1.8	−0.9	4.0	6.0	12.1
Standard error		—	—	4.0	6.9	4.5	6.0	7.4	7.4
Relative standard error[d]		—	—	23%	39%	25%	34%	42%	42%

[a] Henry's law constant H is given in Pa*L/mol, and all energy values are given in kJ/mol; the experimental data are collected from different literature sources.[16, 26–38]

[b] Calculated from experimental vapour pressure P_V and water solubility S_W as P_V/S_W.

[c] P_V value at 25° interpolated from data at lower and higher temperatures.

[d] Calculated as (standard error)/(average experimental ΔG_S).

Results and Discussion

For the 40 compounds listed in Table I, Henry's law constant varies by almost eight orders of magnitude, which corresponds to a variation of experimental ΔG_s of 44 kJ/mol. The greatest affinity to water is observed for hydroquinone ($\log H = -1.50$ with H in Pa*L/mol) followed by *m*-phenylenediamine and 1,3-benzenediol, and great volatilities from water with $\log H$ values above 5 are found for benzene as well as for the methylated and halogenated derivatives.

Interestingly, thiophenol is more volatile from water than phenol by three orders of magnitude, which can be traced back to both a greater vapour pressure and a lower water solubility of the compound.[16] Further striking features of the data set are the very similar $\log H$ values for benzene, toluene, *p*-xylene and fluorobenzene, the distinct variation of Henry's law constant within closely related isomers like 2-chlorotoluene and 3-chlorotoluene, 1,2-benzenediol and 1,3-benzenediol as well as *o*-phenylenediamine and *m*-phenylenediamine, and the very similar $\log H$ values for benzaldehyde, benzonitrile and nitrobenzene that reflect the balance between decreasing vapour pressures (158, 102, and 31.6 Pa) and decreasing water solubilities (64.6, 39.8, and 15.1 mmol/L). Note further the almost identical Henry's law constants for 1,4-dichlorobenzene and bromobenzene, and the great difference between Henry's law constants of *m*-chloronitrobenzene and *m*-bromonitrobenzene, which contrasts with very similar values for chlorobenzene and bromobenzene. The data set thus forms an interesting test case for the sensitivity of the various DCMs to subtle changes in the molecular structure.

OVERALL PERFORMANCE OF MODELS PARAMETERIZED TO PREDICT SOLVATION FREE ENERGIES

Comparison of experimental and calculated ΔG_s values reveals large systematic errors as well as inconsistencies among the individual prediction schemes. As can be seen from Table I, the smallest average error is observed for SM5.0R (−0.9 kJ/mol), the only nonquantum chemical method that treats electrostatics implicitely. SM2 and SM5.4A yield still quite small systematic errors, while MST-AM1, SCF-PCM and MP2-PCM show increasingly greater underestimations of the energy stabilization through aqueous solvation by 4, 6, and 12.1 kJ/mol, respectively. Individual deviations surpass the average errors by far, as is also reflected by the standard errors in the bottom of Table I.

OUTLIERS WITH SM2, SM5.4A, AND SM5.0R

With SM2, the greatest overestimation by −11.7 kJ/mol of the (negative) solvation free energy is observed for *m*-phenylenediamine, and the greatest underestimation is 5.6 kJ/mol for 1,2,3-trimethoxybenzene. Moreover, SM2 shows greater overestimations for the dibromobenzenes and iodobenzene, which contrasts with a significantly better performance for the three chlorobenzenes. The latter might indicate that the SM2 parameterization is inferior for Br and I at aromatic rings compared to aromatic Cl substituents.

Although SM2 reproduces fairly well the difference in ΔG_s between *o*- and *m*-phenylenediamine, SM5.4A underestimates this value by 8.4 kJ/mol. For 2,4,6-trinitrotoluene, SM5.4A yields the greatest underestimation of the energy gain through solvation by 32.1 kJ/mol, followed by 2,4-dinitrotoluene (17.4 kJ/mol) and *m*-bromonitrobenzene (11.5 kJ/mol). The greatest negative deviations are achieved for 1,3-benzenediol (−12.1 kJ/mol), 1,2-benzenediol (−6.9 kJ/mol), and hydroquinone (−5.4 kJ/mol). These findings suggest problems of SM5.4A with multiple NO_2 and OH substituents at aromatic rings.

Interestingly, 2,4,6-trinitrotoluene as compound with the largest positive error of SM5.4A gives the largest negative error with SM5.0R (−14.4 kJ/mol). Further compounds with greater overestimations of ΔG_s are 1,2-benzenediol (−10.6 kJ/mol) and 1,3-benzenediol (−9.0 kJ/mol), and the largest positive errors are seen for *m*-phenylenediamine (6.6 kJ/mol) and *N*,*N*-dimethylamine (6.3 kJ/mol). Note further that SM5.0R fails to reproduce the significant difference in ΔG_s between *o*- and *m*-phenylenediamine, suggesting that this difference is mainly driven by electrostatics.

OUTLIERS WITH MST-AM1 AND PCM-vdW

With MST-AM1, the greatest negative and positive error is −6.9 kJ/mol for 2-nitrotoluene, and 16.4 kJ/mol for hydroquinone. Further large underestimations of the energy gain through aqueous solvation are achieved for phenol (13.6 kJ/mol), all three cresols (11.1 to 11.7 kJ/mol) and the other two benzenediols (9.9 and 11.1 kJ/mol), for the dibromobenzenes (14.2 and 16.8 kJ/mol) and for 2,6-dichlorobenzonitrile (16 kJ/mol). These examples suggest systematic shortcomings of the current

MST-AM1 parameterization with aromatic OH and multiple higher halogens.

The MP2 level of PCM-vdW//6-31G** yields an average decrease in the calculated energy stabilization through aqueous solvation by 6 kJ/mol compared to the SCF level, which probably reflects the fact that the MP2 correction refers to SCF geometries optimized in the gas phase. The largest positive deviations are calculated for 1,2,3-trimethoxybenzene (SCF: 27.8 kJ/mol, MP2: 33.6 kJ/mol) and *N*,*N*-dimethylaniline (SCF: 25.9 kJ/mol, MP2: 29.9 kJ/mol), respectively. Note further the already substantial difference for anisol (SCF: 15.6 kJ/mol, MP2: 19.4 kJ/mol); these findings suggest a systematic problem of the PCM-vdW//6-31G** parameterization with aromatic OCH_3 substituents.

An interesting case is given by the comparison of 2,4-dinitrotoluene and 2,4,6-trinitrotoluene: Despite significant errors in the absolute values of the PCM-vdW solvation-free energies, the experimental difference in ΔG_s of −4.2 kJ/mol is quite well reproduced with MP2 (−3.5 kJ/mol), but significantly overestimated by the SCF level (−11.5 kJ/mol). Moreover, 2,4,6-trinitrotoluene yields the greatest difference in calculated ΔG_s between SCF and MP2 of 26.5 kJ/mol. In comparison to the related semiempirical MST scheme, PCM-vdW yields much more negative ΔG_s values with a better agreement with experiment for the three benzenediols. The comparison between anisol and benzaldehyde provides an opposite example: The agreement with the experiment is better for MST than for PCM-vdW, and the experimental difference in ΔG_s of −4.2 kJ/mol between the two compounds is well reflected by MST (−4.5 kJ/mol), but significantly overestimated by PCM-vdW (SCF: −15.7 kJ/mol, MP2: −9 kJ/mol).

COMPARATIVE PERFORMANCE WITH RESPECT TO LARGE OR SMALL DIFFERENCES IN ΔG_s

The experimental difference in ΔG_s of 17 kJ/mol between phenol and thiophenol is best reproduced by SM5.4A (14.4 kJ/mol) and SM5.0R (14.3 kJ/mol), while SM2 and SCF/PCM-vdW yield significantly lower values (10.8 and 10.4 kJ/mol, respectively), and MP2//PCM-vdW and MST-AM1 are far off (8.6 and 5.4 kJ/mol, respectively). The quite small ΔG_s variation of 0.8 kJ/mol between benzene, toluene, *p*-xylene and fluorobenzene is better reflected by SM2 (2.3 kJ/mol) than by the other five methods. Here, MST (5.5 kJ/mol) gives a clearly more negative solvation energy for *p*-xylene, and both SCF/PCM-vdW (15 kJ/mol) and MP2//PCM-vdW (14.4 kJ/mol) show significant differences in ΔG_s for all four compounds. Interestingly, the very small ΔG_s variation of 0.2 kJ/mol between benzaldehyde, benzonitrile and nitrobenzene is somewhat better reproduced by the semiempirical schemes (SM2: 3.7 kJ/mol, SM5.4A: 6 kJ/mol, MST: 4.8 kJ/mol) than by PCM-vdW (SCF: 6.9 kJ/mol, MP2: 6.1 kJ/mol), while the opposite holds true for the practically identical solvation free energies of 1,4-dichlorobenzene and bromobenzene (SM2: 2.5 kJ/mol, SM5.4A: 2.7 kJ/mol, MST: 5.4 kJ/mol, SCF/PCM-vdW: 1.0 kJ/mol, MP2//PCM-vdW: 1.3 kJ/mol). Note that in both of these cases, the differences in ΔG_s according to SM5.0R come closest to the experimental data (3.3 kJ/mol and 0.2 kJ/mol, respectively).

All six methods fail to mimic the −10 kJ/mol difference in ΔG_s between *m*-chloronitrobenzene and *m*-bromonitrobenzene (SM2: −3.3 kJ/mol, SM5.4A: −1.7 kJ/mol, SM5.0R: −0.5 kJ/mol, MST: 0.3 kJ/mol, SCF/PCM-vdW: 1.2 kJ/mol, MP2//PCM-vdW: 1.1 kJ/mol), which forms an interesting example of a more complex substituent effect (replacement of Cl by Br as second substituent) in a polar molecule. Similarly, each method yields very similar ΔG_s values for the isomeric chlorotoluenes, which contrasts with the experimental difference of 3.8 kJ/mol. The other isomer pairs brenzcatechin (1,2-benzenediol) and resorcin (1,3-benzenediol) as well as *o*-phenylenediamine and *m*-phenylenediamine with experimental ΔG_s differences of −8.6 kJ/mol and −12.9 kJ/mol show again substantial differences in the performances of the individual methods: For the former pair, PCM-vdW (SCF: −9.5 kJ/mol, MP2: −8.4 kJ/mol), MST (−7.4 kJ/mol) and SM5.0R (−7.0 kJ/mol) are superior to SM2 (−5.2 kJ/mol) and SM5.4A (−13.8 kJ/mol), while the latter difference is better reflected by SM2 (−15.1 kJ/mol) and PCM (SCF: −8.4 kJ/mol, MP2: −8 kJ/mol) than by MST (0.4 kJ/mol), SM5.4A (−4.5 kJ/mol) and SM5.0R (4 kJ/mol). Clearly, these examples form suitable test cases for future refinements of the DCM parameters, and the present shortcomings with substituent effects as well as with effects of the structural variation between isomers indicate that care is needed when applying these methods for more complex chemical structures.

ELECTROSTATIC AND NONELECTROSTATIC COMPONENTS TO THE SOLVATION FREE ENERGY

Recall that the five DCMs discussed so far have been parameterized to yield free energies of solvation, which are calculated from electrostatic and nonelectrostatic contributions (however, without explicitely considering vibrational, rotational, and translational components of the entropy term). Thus, a decomposition of the calculated ΔG_s values into their electrostatic and nonelectrostatic components may provide a better insight into possible causes for the observed deviations between theory and experiment.

With SM2 and SM5.4A, the corresponding components are identified by $\Delta G_s^{\text{SM2-ENP}}$ and $\Delta G_s^{\text{SM5.4A-ENP}}$ (change of electronic and nuclear energy of solute plus solute–solute and solute–solvent polarization interaction), and by $\Delta G_s^{\text{SM2-CDS}}$ and $\Delta G_s^{\text{SM5.4A-CDS}}$ (energy for cavity formation, dispersion interaction, and solvent-structural rearrangement). For MST-AM1 and PCM-vdW//6-31G**, the notations $\Delta G_s^{\text{MST-el}}$, $\Delta G_s^{\text{SCF-el}}$ and $\Delta G_s^{\text{MP2-el}}$ (electrostatic components) as well as $\Delta G_s^{\text{MST-nel}}$ (summarizing the energies of cavitation and van der Waals interaction) and $\Delta G_s^{\text{SCF-nel}}$ (cavitation energy plus dispersion and repulsion inter action) are used.

The corresponding data are summarized in Table II together with results from the COSMO method, which yields electrostatic components in terms of semiempirical solvation enthalpies $\Delta H_s^{\text{COSMO}}$ (COSMO-AM1) as well as DFT solvation energies ΔE_s^{BPW} (COSMO-DFT/BPW-dnp) and ΔE_s^{BLYP} (COSMO-DFT/BLYP-dnp), respectively.

PCM-vdW AND MST-AM1 ELECTROSTATICS

Comparison of the PCM-vdW error statistics of Tables I and II shows that restriction to the electrostatic component yields significantly smaller standard errors for both the SCF and MP2 level. In particular, $\Delta G_s^{\text{MP2-el}}$ comes quite close to the experimental ΔG_s values (average error = −2.8 kJ/mol), and shows the overall best performance of all methods under analysis. With the related semiempirical MST scheme, restriction to the electrostatic component causes a downward shift of the calculated solvation energy by 8.5 kJ/mol on the average, and both average error and standard error are a bit greater than for ΔG_s^{MST}.

COSMO-AM1 AND COSMO-DFT

The greatest standard error with respect to the experimental solvation free energies is achieved by COSMO-AM1. However, inspection of the data distribution reveals systematic calculation errors for benzene derivatives containing the nitro group; exclusion of the respective six compounds reduces the standard error from 16.5 to 4.9 kJ/mol. In this context it should be noted that in contrast to other semiempirical DCMs, the parameterization of COSMO-AM1 has not yet been optimized.

The performance of COSMO-DFT (using atomic radii optimized for a DFT implementation of COSMO-RS)[24] is much better than COSMO-AM1, and (regarding relative trends) comparable to PCM-vdW, yielding standard errors of 3.8 kJ/mol (BPW) and 4.2 kJ/mol (BLYP), respectively. The resultant data distribution is illustrated in Figure 1, where ΔG_s^{MP2}, $\Delta G_s^{\text{MP2-el}}$, and ΔE_s^{BPW} are plotted against experimental free energies of solvation for all 40 compounds. As can be seen from the figure, ΔG_s^{MP2} (which includes nonelectrostatic terms) yields a much greater scatter than $\Delta G_s^{\text{MP2-el}}$ and a systematic underestimation of the energy stabilization through aqueous solvation (cf. Table I), while $\Delta G_s^{\text{MP2-el}}$ provides a good relative trend with systematically too

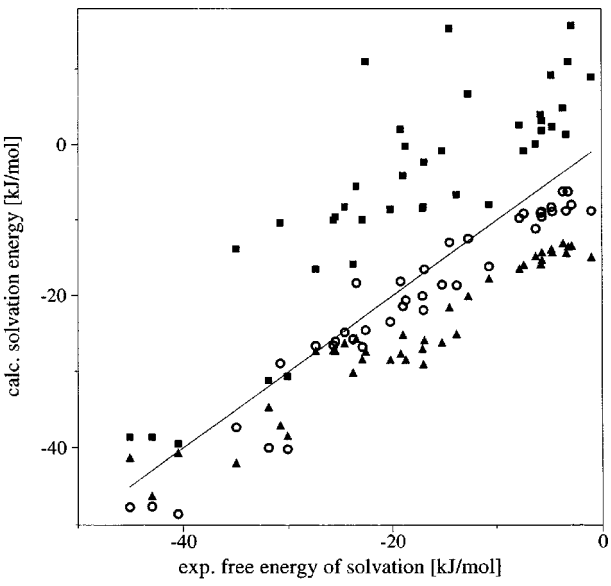

FIGURE 1. Calculated solvation energies according to ΔG_s^{MP2} (MP2//PCM-vdW//6-31G**, filled squares), $\Delta G_s^{\text{MP2-el}}$ (electrostatic component of ΔG_s^{MP2}, circles) and ΔE_s^{BPW} (COSMO-DFT/BPW-dnp, filled triangles) against experimental ΔG_s values for all 40 compounds (cf. Tables I and II). The diagonal is shown as a line.

TABLE II.

Electrostatic and Nonelectrostatic Contributions to the Free Energy of Solvation.

	Electrostatic Contribution (kJ/mol)								Nonelectrostatic Contribution (kJ/mol)			
	DCM-AM1				PCM-vdW//6-31G**		COSMO-DFT/dnp		DCM-AM1			PCM-vdW//6-31G**
No.	$\Delta G_S^{\text{SM2-ENP}}$	$\Delta G_S^{\text{SM5.4A-ENP}}$	$\Delta G_S^{\text{MST-el}}$	$\Delta H_S^{\text{COSMO}}$	$\Delta G_S^{\text{SCF-el}}$	$\Delta G_S^{\text{MP2-el}}$	ΔE_S^{BPW}	ΔE_S^{BLYP}	$\Delta G_S^{\text{SM2-CDS}}$	$\Delta G_S^{\text{SM5.4A-CDS}}$	$\Delta G_S^{\text{MST-nel}}$	$\Delta G_S^{\text{SCF-nel}}$
1	−8.2	−13.0	−8.8	−14.2	−9.1	−6.2	−13.0	−11.9	6.0	8.9	3.8	11.1
2	−8.1	−13.2	−10.1	−14.7	−9.0	−6.2	−13.4	−12.0	6.9	9.5	4.4	17.1
3	−10.5	−28.3	−19.8	−32.7	−30.8	−26.6	−27.2	−26.1	−13.7	0.9	6.0	10.1
4	−8.5	−11.6	−8.1	−15.9	−12.0	−8.8	−14.2	−13.6	3.7	6.7	6.7	11.2
5	−9.2	−11.8	−8.5	−17.5	−15.1	−11.1	−14.7	−14.4	2.1	4.9	9.1	11.2
6	−7.2	−8.7	−6.6	−16.6	−12.6	−9.5	−14.2	−14.0	1.5	4.5	9.5	11.4
7	−9.8	−11.7	−9.2	−18.5	−12.2	−8.8	−15.2	−14.4	1.5	4.8	7.6	12.0
8	−12.6	−12.5	−10.5	−22.9	−13.0	−9.7	−16.4	−16.0	−1.7	1.7	10.4	12.3
9	−10.3	−8.7	−7.8	−21.5	−12.1	−9.0	−15.8	−15.5	−3.1	0.7	11.2	13.0
10	−14.9	−18.8	−20.9	−31.3	−29.1	−25.7	−30.1	−28.0	−13.4	−1.0	5.5	9.9
11	−12.0	−18.4	−16.4	−27.6	−16.2	−12.4	−20.0	−19.1	2.2	5.6	6.2	19.1
12	−17.0	−25.4	−21.6	−41.8	−27.0	−16.5	−25.8	−25.7	−2.7	7.0	6.9	14.1
13	−10.0	−28.2	−18.9	−30.1	−28.6	−24.8	−26.2	−24.8	−10.5	5.4	6.0	16.6
14	−10.5	−28.7	−21.0	−33.2	−30.2	−26.0	−27.2	−26.0	−12.8	1.5	6.6	16.5
15	−10.3	−28.2	−20.9	−32.9	−30.4	−26.5	−27.1	−25.8	−12.8	1.5	6.6	16.5
16	−13.2	−36.2	−28.7	−54.8	−31.3	−20.0	−26.9	−27.6	−8.7	23.4	11.6	11.5
17	−13.5	−14.6	−15.6	−29.0	−18.5	−16.1	−17.7	−16.5	0.1	1.6	7.2	8.1
18	−7.9	−13.3	−14.1	−15.1	−11.4	−7.9	−13.3	−11.7	7.9	10.0	4.9	23.7
19	−12.4	−43.1	−36.5	−51.1	−53.4	−47.7	−41.3	−40.3	−33.4	−7.4	7.8	9.1
20	−11.3	−21.4	−19.7	−28.4	−29.3	−21.9	−29.0	−28.6	−6.8	9.0	7.3	13.7
21	−16.2	−25.4	−31.0	−43.2	−30.3	−24.5	−27.3	−26.8	−0.9	7.0	9.0	35.6
22	−9.5	−13.0	−10.9	−16.2	−12.8	−8.7	−14.3	−13.8	9.2	12.1	7.2	10.1
23	−8.6	−12.0	−10.2	−15.6	−11.5	−8.2	−13.8	−13.0	4.7	7.7	6.8	17.3
24	−8.4	−11.6	−10.7	−16.0	−12.0	−8.7	−14.8	−14.0	4.7	7.3	7.1	17.7
25	−13.0	−10.7	−10.9	−17.2	−12.9	−9.1	−15.9	−14.9	−1.3	3.9	8.1	8.3
26	−11.9	−34.7	−29.8	−42.6	−44.8	−40.0	−34.6	−33.5	−28.4	−4.1	7.8	8.8
27	−12.1	−45.8	−37.3	−51.3	−54.5	−48.6	−40.7	−39.7	−33.3	−6.8	7.9	9.1
28	−11.4	−21.1	−40.6	−53.8	−44.6	−40.2	−38.4	−36.7	−27.6	−9.7	6.2	9.5
29	−21.5	−24.5	−40.5	−50.3	−52.4	−47.6	−46.2	−43.3	−32.6	−10.8	6.4	8.9
30	−11.2	−20.8	−20.7	−28.1	−27.8	−20.6	−28.4	−28.0	−5.1	9.0	7.9	20.4

TABLE II.
(Continued)

	Electrostatic Contribution (kJ/mol)								Nonelectrostatic Contribution (kJ/mol)			
	DCM-AM1				PCM-vdW//6-31G**		COSMO-DFT/dnp		DCM-AM1			PCM-vdW//6-31G**
No.	$\Delta G_s^{\text{SM2-ENP}}$	$\Delta G_s^{\text{SM5.4A-ENP}}$	$\Delta G_s^{\text{MST-el}}$	$\Delta H_s^{\text{COSMO}}$	$\Delta G_s^{\text{SCF-el}}$	$\Delta G_s^{\text{MP2-el}}$	ΔE_s^{BPW}	ΔE_s^{BLYP}	$\Delta G_s^{\text{SM2-CDS}}$	$\Delta G_s^{\text{SM5.4A-CDS}}$	$\Delta G_s^{\text{MST-nel}}$	$\Delta G_s^{\text{SCF-nel}}$
31	−12.5	−21.2	−17.0	−28.9	−30.7	−23.4	−28.4	−28.9	−7.9	4.4	12.8	14.9
32	−10.6	−33.9	−33.0	−51.2	−28.6	−18.5	−26.1	−26.8	−6.8	21.8	10.9	17.7
33	−9.5	−49.7	−48.2	−91.5	−47.4	−28.9	−37.0	−39.3	−21.3	36.4	18.1	18.6
34	−1.0	−52.9	−58.5	−126.4	−63.8	−37.3	−42.0	−45.3	−34.7	50.0	25.0	23.6
35	−10.5	−31.5	−29.4	−54.7	−29.2	−18.6	−25.0	−26.1	−10.9	21.2	13.9	12.0
36	−11.5	−31.4	−30.0	−57.6	−28.8	−18.3	−25.6	−26.7	−13.2	19.3	14.6	12.8
37	−17.9	−26.2	−28.4	−40.4	−28.4	−18.1	−27.6	−27.2	−0.6	6.7	6.9	20.1
38	−10.4	−19.4	−28.3	−36.8	−30.8	−26.7	−28.3	−26.5	−10.5	0.8	5.7	16.7
39	−15.9	−18.0	−25.2	−32.1	−25.6	−21.4	−25.1	−23.4	−2.8	0.1	5.3	17.3
40	−18.5	−17.9	−22.5	−28.4	−16.9	−12.9	−21.5	−19.7	5.8	4.3	5.8	28.4
Average	−11.5	−22.9	−22.2	−35.8	−26.6	−20.5	−24.7	−24.1	−7.5	7.0	8.5	14.9
Range of variation	20.5	44.2	51.9	112.2	54.8	42.4	33.2	33.6	43.9	60.8	21.2	27.5
Average error	6.3	−5.2	−4.4	−18.1[b]	−8.9	−2.8	−7.0	−6.4	10.2	24.7	26.2	32.6
Standard error	11.4	7.4	7.1	16.5[b]	4.8	3.0	3.8	4.2	4.7	15.5	13.3	12.9
Relative standard error[a]	64%	42%	40%	93%[b]	27%	17%	21%	24%	27%	88%	75%	73%

[a] Calculated as (standard error)/(average experimental ΔG_s).
[b] For the subset of 34 compounds without NO_2 congeners, the following statistics are obtained: average error = −12.4 kJ/mol, standard error = 4.9 kJ/mol, relative standard error = 29%.

negative solvation energies, the latter of which is slightly more pronounced with COSMO-DFT.

ELECTROSTATIC COMPONENTS AND CDS TERMS

Both SM2-ENP and SM2-CDS are significantly inferior to SM2 in reproducing the experimental solvation free energies. The standard error of $\Delta G_s^{\text{SM2-ENP}}$ is more than twice the one of $\Delta G_s^{\text{SM2-CDS}}$, and correlation analysis reveals the following suprising results: While the electrostatic PCM component $\Delta G_s^{\text{MP2-el}}$ shows relatively high squared correlation coefficients with the (electrostatic) COSMO energies ΔE_s^{BPW} ($r^2 = 0.92$) and ΔE_s^{BLYP} ($r^2 = 0.87$) as well as with ΔG_s^{SM2} ($r^2 = 0.92$), there is almost no correlation with $\Delta G_s^{\text{SM2-ENP}}$ ($r^2 = 0.05$), and an only moderate correlation with $\Delta G_s^{\text{SM5.4A-ENP}}$ ($r^2 = 0.55$). Note further the surprisingly high correlation of $\Delta G_s^{\text{MP2-el}}$ with $\Delta G_s^{\text{SM2-CDS}}$ ($r^2 = 0.90$), which contrasts with the corresponding value for $\Delta G_s^{\text{SM5.4A-CDS}}$ ($r^2 = 0.03$). The individual trends of $\Delta G_s^{\text{SM2-ENP}}$ and $\Delta G_s^{\text{SM5.4A-ENP}}$ as well as of $\Delta G_s^{\text{SM2-CDS}}$ and $\Delta G_s^{\text{SM5.4A-CDS}}$ have practically nothing in common (r^2 values of 0.002 and 0.001, respectively).

These findings suggest that the ENP component of the SM2 scheme encodes solvation energy contributions very different from the electrostatic components of all other schemes, while the CDS component contains information that in its relative trend is much more similar to electrostatic solvation components than would be expected from its definition to contain cavitation and dispersion energies. It appears further that despite a significantly lower correlation of ΔG_s^{exp} with $\Delta G_s^{\text{SM5.4A}}$ than with ΔG_s^{SM2} ($r^2 = 0.91$ vs. $r^2 = 0.70$), the SM5.4A partitioning into ENP and CDS components is somewhat more realistic than the SM2 partitioning, in accord with a corresponding discussion upon introduction of the SM5.4A model.[21]

COMPARISON OF NONELECTROSTATIC COMPONENTS

Comparative analysis of the differently calculated nonelectrostatic components reveals further interesting features: only $\Delta G_s^{\text{SM2-CDS}}$ shows a significant correlation with the experimental solvation free energies ($r^2 = 0.86$), and relative trends of all four energy terms have very little in common as indicated by r^2 values of 0.06 (SM2-CDS vs. SCF-nel), (−)0.15 (SM2-CDS vs. MST-nel), 0.15 (SM2-CDS vs. SCF-nel), 0.13 (SM5.4A-CDS vs. MST-nel), 0.60 (SM5.4A-CDS vs. MST-nel), and 0.02 (SCF-nel vs. MST-nel). The respective data distributions compared with the experimental solvation-free energies are plotted in Figure 2, showing again the significant correlation of $\Delta G_s^{\text{SM2-CDS}}$ with ΔG_s, which contrasts with the trends of $\Delta G_s^{\text{MST-nel}}$ and $\Delta G_s^{\text{SCF-nel}}$. It follows that with all schemes, the nonelectrostatic component distinctly depends on the overall parameterization of the respective DCM, and that a separate use of any of these energy terms is not recommended without further analysis.

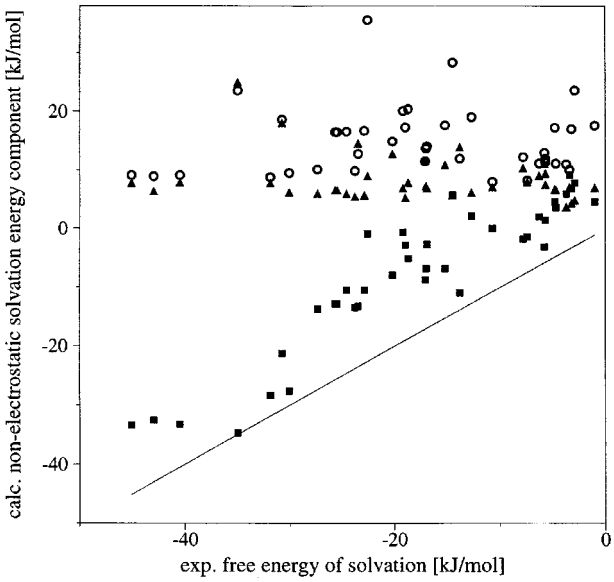

FIGURE 2. Calculated nonlectrostatic components of the solvation energy according to $\Delta G_s^{\text{SM2-CDS}}$ (SM2-AM1, filled squares), $\Delta G_s^{\text{SCF-nel}}$ (PCM-vdW// 6-31G** circles), and $\Delta G_s^{\text{MST-nel}}$ (MST-AM1, filled triangles) against experimental ΔG_s values for all 40 compounds (cf. Tables I and II). The diagonal is shown as a line.

Inspection of Table II shows further, that $\Delta G_s^{\text{SM2-ENP}}$ and $\Delta G_s^{\text{SM5.4A-ENP}}$ yield negative energy values for all compounds, and in this respect agree with the electrostatic solvation components of MST-AM1, PCM-vdW//6-31G**, COSMO-AM1, and COSMO-DFT. On the other hand, with $\Delta G_s^{\text{SM2-CDS}}$ and $\Delta G_s^{\text{SM5.4A-CDS}}$ both positive and negative solvation components are achieved, while $\Delta G_s^{\text{MST-nel}}$ and $\Delta G_s^{\text{SCF-nel}}$ yield destabilizing contributions to the solvation free energy for all compounds of the test set. For phenol, all cresols, benzenediols and phenylenediamines, di- and trinitrotoluene, *m*-chloronitrobenzene and *m*-bromonitrobenzene, and 2-methylaniline the solvation contribution from $\Delta G_s^{\text{SM2-CDS}}$ is even more negative than from $\Delta G_s^{\text{SM2-ENP}}$. Among these 13 compounds, 2,4,6-trinitrotoluene is the molecule with the great-

est negative solvation contribution from $\Delta G_s^{\text{SM2-CDS}}$ (−34.7 kJ/mol), and with the greatest difference between the energy components calculated by $\Delta G_s^{\text{SM2-CDS}}$ and by $\Delta G_s^{\text{SM2-ENP}}$ (−33.7 kJ/mol). Note further that the total variation of $\Delta G_s^{\text{SM2-CDS}}$ across the data set is more than twice as large as the one of $\Delta G_s^{\text{SM2-ENP}}$ (43.9 vs. 20.5 kJ/mol).

SM5.4A yields negative CDS contributions only for the three benzenediols and two phenylenediamines, and all of these values are smaller in absolute size than the ones from SM2-CDS and from SM5.4A-ENP. It is unclear whether these negative SM5.4A-CDS values represent electrostatic portions of the first solvation-shell contributions to ΔG_s,[21] or whether the overall parameterization scheme has lead to an implicit account of entropic contributions in the solvent-structural rearrangement portion of the CDS component.

GAS-PHASE DIPOLE MOMENTS AND ELECTROSTATIC SOLVATION ENERGIES

Although the sum of higher-order moments may well exceed the dipole moment regarding the electrostatic contribution to the solvation energy, calculated dipole moments are still instructive for characterizing differences in the electronic structure description of the different continuum-solvation schemes. Moreover, compounds with large dipole moments are expected to have large negative solvation energies, and in such cases significant errors in the calculated dipole moments would lead to corresponding deviations between calculated and experimental ΔG_s values.

As can be seen from Table III, HF/6-31G** yields greater gas-phase dipole moments on the average than AM1, MP2//6-31G**, and DFT. Chlorinated and brominated derivatives are examples where, except for 1,2-dichlorobenzene, the electrostatic component of PCM-vdW is less stabilizing than COSMO-DFT, although the HF/6-31G** dipole moments are greater by ca. 0.5 Debye than with DFT.

For aromatic rings with NO_2 substituents (Nos. 16, 32–36), the results change significantly upon inclusion of electron correlation: HF/6-31G** yields dipole moments greater than MP2//6-31G** and DFT/BPW-dnp by around 1 Debye and 0.5 Debye, respectively, and the electrostatic solvation energies decrease (in absolute size) correspondingly from $\Delta G_s^{\text{SCF-el}}$ to ΔE_s^{BPW} to $\Delta G_s^{\text{MP2-el}}$. With *m*- and *p*-benzenediol (Nos. 26–27), both dipole moment and electrostatic solvation energy decrease in the order HF, MP2, DFT.

Note further that *N*,*N*-dimethylaniline is the compound with the greatest difference in gas-phase dipole moments between HF/6-31G** and DFT/BPW-dnp (1.5 Debye), which is mainly caused by different geometries: With HF/6-31G**, the nitrogen of the dimethylamino group is pyramidal (in agreement with AM1 and COSMO-AM1), while DFT/BPW-dnp (like COSMO-DFT/BPW-dnp) predicts a planar configuration with both methyl carbons almost in plane with the aromatic ring. Here, the difference between ΔE_s^{BPW} and $\Delta G_s^{\text{SCF-el}}$ of −4.6 kJ/mol is smaller than for some nitrobenzenes and benzenediols. With 2,4,6-trinitrotoluene, the much greater electrostatic solvation energy of PCM-vdW compared to COSMO-DFT is clearly not driven by the dipole moment contribution.

GAS-PHASE VS. SOLUTION-PHASE DIPOLE MOMENTS

In accord with expectation the solution-phase dipole moments are generally higher than the ones in the gas phase, with an average difference of around 0.8 Debye for all schemes except SM2 and SM5.4A, which yield average differences to AM1 of 0.5 and 1.3 Debye, respectively (note that the gas-phase dipole moments derived from CM1A charges[53] that are used for the SM5.4A model are larger by 0.1 Debye on the average than with AM1). The polarizing effect of aqueous solution is further reflected by significant correlations between the solution-phase increase in dipole moment and the absolute size of the gas-phase dipole moment: The respective r^2 values are 0.99 (COSMO-DFT/BPW-dnp), 0.91 (SCF/PCM-vdW, ICOMP = 4), 0.83 (SCF/PCM-vdW, ICOMP = 2), 0.86 (MP2//PCM-vdW, ICOMP = 2), 0.81 (SM2), 0.83 (SM5.4A), 0.88 (MST-AM1), and 0.93 (COSMO-AM1). Moreover, COSMO-AM1 yields particularly large solution-phase dipole moments for the six benzene derivatives containing NO_2, which is in line with the observed substantial overestimation of the (negative) solvation energy by $\Delta H_s^{\text{COSMO}}$ as mentioned above.

SM2 and SM5.4A are the only methods where in a few cases aqueous solvation is predicted to lower the dipole moment of the compound. With SM2, these compounds are thiophenol, iodobenzene, and *m*-phenylenediamine. With SM5.4A, the apparently too small solution-phase dipole moments of aniline, *o*- and *m*-phenylenediamine (Nos. 10, 28–29) can be traced back to the underlying CM1A charges,[21, 53] which yield much smaller gas-phase dipole moments than standard AM1 (0.7 vs. 1.6 Debye, 0.1

TABLE III.
Gas-Phase and Solution-Phase Dipole Moments in Debye.

	Gas Phase				Aqueous Phase						
No.	AM1	HF/6-31G**	MP2// 6-31G**	DFT/ BPW-dnp	SM2	SM5.4A	MST-AM1	PCM-vdW// 6-31G** ICOMP=4[a]	PCM-vdW// 6-31G** ICOMP=2[a]	MP2//PCM-vdW// 6-31G** ICOMP=2[a]	COSMO-DFT/ BPW-dnp
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.3	0.3	0.3	0.5	0.3	0.4	0.4	0.4	0.4	0.4	0.6
3	1.2	1.4	1.4	1.2	1.5	2.2	1.8	2.0	2.0	1.9	1.7
4	1.3	2.1	1.8	1.6	1.9	2.5	1.9	3.0	3.0	2.8	2.3
5	2.0	3.2	2.7	2.3	2.8	4.0	2.9	4.6	4.6	4.1	3.3
6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	1.4	2.1	1.9	1.6	1.7	2.5	2.1	2.7	3.3	3.1	2.3
8	2.1	3.0	2.6	2.2	2.5	3.8	3.1	3.8	4.7	4.4	3.1
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	1.6	1.6	1.8	2.0	1.7	1.1	2.1	2.2	2.2	2.4	2.8
11	1.2	1.4	1.3	1.3	1.5	1.8	1.7	1.8	1.8	1.7	1.8
12	2.9	3.5	2.9	3.5	4.1	4.8	4.2	4.7	4.7	4.1	4.9
13	1.0	1.2	1.2	0.9	1.1	1.8	1.3	1.6	1.6	1.6	1.2
14	1.1	1.2	1.1	0.8	1.3	2.0	1.5	1.6	1.6	1.6	1.1
15	1.3	1.4	1.3	1.2	1.6	2.3	1.9	2.0	2.0	1.8	1.7
16	5.2	5.1	4.1	4.6	6.8	8.8	6.9	6.6	6.6	5.7	6.4
17	1.2	1.9	1.8	1.6	0.7	1.5	1.8	2.5	2.5	2.4	2.2
18	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
19	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
20	3.3	4.9	4.3	4.6	4.5	6.4	5.1	6.2	6.2	5.7	6.3
21	1.8	1.8	1.6	1.7	2.2	2.9	2.6	2.4	2.5	2.3	2.4
22	1.6	1.7	1.3	1.5	2.3	2.7	2.4	2.1	2.4	2.0	2.1
23	1.1	1.9	1.6	1.4	1.6	2.3	1.8	2.5	2.8	2.5	2.0
24	1.5	2.4	2.1	2.0	2.0	2.7	2.3	2.9	3.3	3.0	2.7
25	1.4	1.9	1.7	1.6	1.3	2.3	2.0	2.3	3.2	3.0	2.2
26	2.1	2.6	2.5	2.2	2.6	3.8	3.2	3.5	3.5	3.4	3.1
27	2.4	2.7	2.3	2.1	3.0	4.4	3.6	3.6	3.7	3.3	2.9
28	0.4	0.7	1.0	1.2	0.7	0.0	0.5	1.1	1.2	1.5	1.9
29	1.5	1.6	1.7	1.8	1.1	1.1	2.1	2.1	2.2	2.4	2.6
30	3.2	4.6	4.1	4.3	4.3	6.4	5.0	5.9	6.0	5.5	6.0

TABLE III.
(Continued)

	Gas Phase				Aqueous Phase						
No.	AM1	HF/6-31G**	MP2// 6-31G**	DFT/ BPW-dnp	SM2	SM5.4A	MST-AM1	PCM-vdW// 6-31G** ICOMP=4[a]	PCM-vdW// 6-31G** ICOMP=2[a]	MP2//PCM-vdW// 6-31G** ICOMP=2[a]	COSMO-DFT/ BPW-dnp
31	4.1	6.1	5.3	5.3	5.6	8.5	6.1	8.2	8.3	7.5	7.6
32	5.0	4.8	3.8	4.3	6.6	8.7	7.1	6.0	6.3	5.4	6.2
33	5.4	5.4	4.3	5.0	7.3	9.5	7.4	6.9	7.0	6.0	7.0
34	1.4	1.6	1.2	1.7	1.6	2.0	1.8	1.8	1.9	1.5	2.3
35	4.7	4.3	3.4	3.9	6.1	7.7	6.3	5.5	5.6	4.8	5.5
36	4.6	4.3	3.4	3.9	6.1	7.7	6.3	5.5	5.5	4.7	5.5
37	3.0	3.3	2.7	3.1	4.2	5.3	4.8	4.5	4.6	4.0	4.5
38	1.6	1.5	1.6	1.7	1.7	1.6	1.8	2.0	2.0	2.1	2.4
39	1.5	1.5	1.6	1.9	1.8	1.5	1.7	1.9	1.9	2.0	2.7
40	1.1	0.9	0.9	2.4	1.5	1.4	1.4	1.2	1.1	1.1	3.4
Average	1.9	2.2	2.0	2.1	2.4	3.2	2.7	2.9	3.1	2.8	2.9
Range of variation	5.4	6.1	5.3	5.3	7.3	9.5	7.4	8.2	8.3	7.5	7.6

[a] ICOMP = 4 denotes the more sophisticated normalization procedure that uses additional surface charges distributed according to the solute electronic density, while with ICOMP = 2 a constant factor is used for the surface charge normalization.[43] For MP2 dipole moments with PCM-vdW, only ICOMP = 2 is available.

vs. 0.4 Debye and 0.7 vs. 1.5 Debye for the three compounds using CM1A and AM1, respectively). Here, *m*-phenylenediamine is the only compound where the (already very small) CM1A gas-phase dipole moment is still reduced in the SM5.4A model. Note, however, that according to gas-phase dipole moments from MP2//6-31G** (1.8, 1.0, 1.7 Debye) and DFT/BPW-dnp (2.0, 1.2, 1.8 Debye) the respective CM1A-derived dipole moments appear indeed to be too small.

With PCM-vdW, the simpler polarization charge normalization with a constant factor (ICOMP = 2)[45] results in significantly larger dipole moments for 1,2-dibromobenzene and iodobenzene compared to the standard approach (ICOMP = 4),[45] which, however, is only available for dipole moments at the HF level. As a consequence, MP2//6-31G** (ICOMP = 2) yields greater dipole moments for these two compounds than HF/6-31G**, while generally the MP2 level yields a lowering of the dipole moments compared to the HF level both in the gas phase and in aqueous solution.

Moreover, the significantly overestimated stabilization through aqueous solvation of nitrobenzene and its derivatives by COSMO-AM1 can be partly traced back to corresponding overestimations of the dipole moments: For the six nitro compounds, the AM1 dipole moments are greater by 0.5 Debye on the average than the ones from DFT/BPW-dnp, and this average difference increases to 0.8 Debye for COSMO-AM1 vs. COSMO-DFT.

Interestingly, the CN nitrogen gives a different picture: AM1 dipole moments of benzonitrile, *o*-toluenenitrile, and 2,6-dichlorobenzonitrile (Nos. 20, 30–31) are lower by 1.1–1.3 Debye than with DFT, and in solution the differences are still slightly larger (4.9, 4.8, and 6.3 Debye vs. 6.3, 6.0, and 7.6 Debye). Nonetheless, COSMO-AM1 and COSMO-DFT/BPW-dnp agree almost perfectly in their solvation energy predictions for these three compounds, showing differences of only −0.6 kJ/mol (No. 20), −0.3 kJ/mol (No. 30), and 0.5 kJ/mol (No. 31), respectively.

An extreme case of opposing trends is given by *o*-phenylenediamine (No. 28): Although the semiempirical gas-phase and solution-phase dipole moments (AM1 and COSMO-AM1) are significantly smaller than the DFT counterparts (0.4 vs. 1.2 Debye and 0.6 vs. 1.9 Debye, respectively), the solvation energy predicted by COSMO-DFT is 15.4 kJ/mol less negative than the one by COSMO-AM1.

ESCAPED CHARGES OF PCM-vdW

Coming back to the PCM-vdW treatment of escaped charges, it is interesting to note that with the current data set, both normalization procedures yield in fact almost identical results in the range of 0.25 to 0.45 atomic units (a.u.). The four largest (and for ICOMP = 4 and ICOMP = 2 almost identical) values are found for 2,6-dichlorobenzonitrile (0.45 a.u.), *N*-methylaniline (0.44 a.u.), anisol (0.42 a.u.), and 1,2,3-trimethoxybenzene (0.41 a.u.), while 1,2-dibromobenzene yields a smaller value (0.28 a.u. with ICOMP = 4, 0.27 a.u. with ICOMP = 2), and with iodobenzene it is somewhat surprising that the escaped charge due to ICOMP = 2 is significantly smaller than with ICOMP = 4 (0.41 a.u. with ICOMP = 4, 0.27 a.u. with ICOMP = 2). However, the latter two compounds show unusually large differences between ICOMP = 4 and ICOMP = 2 regarding the errors of the nuclear polarization charges (1,2-dibromobenzene: 0.19 a.u. vs. 0.61 a.u., iodobenzene: 0.34 a.u. vs. 0.42) and of the electronic polarization charges (1,2-dibromobenzene: 0.43 a.u. vs. −0.85 a.u., iodobenzene: −0.61 vs. −0.68), which reflects the better local charge compensation as achieved by the more sophisticated approach, and may explain the above-mentioned significant differences between the associated dipole moments. Only the other two brominated congeners (Nos. 7 and 9) also show somewhat larger differences in the individual polarization charge errors, while for all other compounds these errors agree within 0.01 a.u. for ICOMP = 4 and ICOMP = 2.

LINEAR REGRESSION MODELS FOR HENRY'S LAW CONSTANT

In Table IV, linear regression results for Henry's law constant are summarized for the various quantum chemical schemes under analysis. Considerable variation is observed both for the statistics as well as for slope and intercept. In agreement with the previous discussion of the results of Table II, the overall best performance is provided by the electrostatic component of MP2//PCM-vdW, where both slope and intercept come relatively close to the theoretical values as mentioned above. Here, the standard error of 0.52 log units corresponds to a factor of 3.3, which is certainly above the (unknown) experimental error.

COSMO-DFT yields the second-best statistics, and the respective slope and intercept indicate that there is a need for scaling $\Delta E_s^{\mathrm{BPW}}$ and $\Delta E_s^{\mathrm{BLYP}}$ when employing the currently used set of parameters. Interestingly, restriction to the gas-phase geometries

TABLE IV.
Linear Regression Results for Henry's Law Constant.[a]

Solvation Energy Parameter	r^2_{adj}	SE	Coefficient	Constant	$F_{1,38}$ ($F_{1,34}$)[b]
ΔG_s^{SM2}	0.91	0.63	0.150 (±0.008)	6.13 (±0.18)	380.9
$\Delta G_s^{SM5.4A}$	0.70	1.14	0.140 (±0.015)	5.52 (±0.30)	90.7
$\Delta G_s^{SM5.0R}$	0.91	0.62	0.140 (±0.007)	5.89 (±0.16)	395.7
ΔG_s^{MST}	0.73	1.07	0.168 (±0.016)	5.59 (±0.28)	109.0
$\Delta G_s^{MST\text{-}el}$	0.67	1.18	0.140 (±0.015)	6.39 (±0.39)	81.9
ΔH_s^{COSMO}	0.45	1.53	0.064 (±0.011)	5.57 (±0.47)	33.0
ΔH_s^{COSMO} ($n = 34$)[b]	0.84	0.86	0.163 (±0.012)	8.20 (±0.39)	173.6
ΔG_s^{SCF}	0.79	0.96	0.118 (±0.010)	4.67 (±0.19)	143.3
ΔG_s^{MP2}	0.72	1.09	0.126 (±0.012)	3.99 (±0.19)	101.4
$\Delta G_s^{SCF\text{-}el}$	0.90	0.67	0.138 (±0.008)	6.96 (±0.23)	334.6
$\Delta G_s^{MP2\text{-}el}$	0.94	0.52	0.167 (±0.007)	6.71 (±0.16)	583.2
ΔE_s^{BPW} (gas-phase geometry)	0.93	0.54	0.220 (±0.010)	8.67 (±0.25)	535.8
ΔE_s^{BPW}	0.93	0.56	0.217 (±0.010)	8.65 (±0.26)	499.0
ΔE_s^{BLYP} (gas-phase geometry)	0.90	0.66	0.212 (±0.011)	8.33 (±0.29)	350.6
ΔE_s^{BLYP}	0.89	0.68	0.208 (±0.011)	8.30 (±0.30)	320.3

[a] Coefficient and constant represent the slope and intercept of the linear regression of log H on the solvation energy parameter, using all 40 compounds as listed in Table I (however, see [b]). The statistical parameters are as follows: r^2_{adj} = squared correlation coefficient, adjusted for degrees of freedom, SE = standard error, F = Fisher test value.

[b] The subset of 34 compounds is defined by omitting nitrobenzene and all five nitrobenzene derivatives from the compound set.

yields very similar and even slightly better results than when performing solution-phase geometry optimization. Note further that the relative trend with COSMO-DFT is superior to all schemes that evaluate both electrostatic and nonelectrostatic terms, including PCM-vdW on the HF and MP2 level.

The MST-AM1 method yields only moderate statistics, which are still somewhat lower when omitting the nonelectrostatic component (explained variances of 73 and 67%, respectively). However, $\Delta G_s^{MST\text{-}el}$ shows a significantly improved performance for the subset of 34 compounds without NO_2 substituents ($r^2_{adj} = 0.81$, SE = 0.93), which is now also better than the one with ΔG_s^{MST} ($r^2_{adj} = 0.75$, SE = 1.07; statistics not shown in Table IV). These results indicate that MST predictions of ΔG_s are less suitable for nitrobenzene derivatives, mainly because of corresponding AM1 deficiences as already discussed with COSMO-AM1.

A further interesting aspect is the fact that for the subset of 34 compounds without NO_2, the HF level of the electrostatic component of PCM-vdW is no longer inferior to the corresponding MP2 level ($r^2_{adj} = 0.95$ in both cases, statistics not shown in Table IV). It suggests that the MP2 correction is particularly important for aromatic nitro compounds; indeed, the MP2 correction to the dipole moment was relatively large for the NO_2-substituted compounds compared to the other benzene derivatives (cf. Table III).

Among the SMx models, SM2 and the nonquantum chemical SM5.0R are clearly superior to SM5.4A when taking all 40 compounds. Omission of the six nitroaromatic compounds gives almost identical statistics for SM2 ($r^2_{adj} = 0.91$, SE = 0.64) and a slightly improved performance of SM5.0R ($r^2_{adj} = 0.92$, SE = 0.59), but a significant improvement for SM5.4A ($r^2_{adj} = 0.92$, SE = 0.62). Interestingly, SM5.4A yields a large underestimation of the stabilizing solvation energy for this subset of nitrobenzene and its derivatives (average error = 12 kJ/mol, standard error = 11.3 kJ/mol), although the gas-phase CM1A dipole moments are larger by around 0.4 Debye than with standard AM1. Apart from that, it should be noted that the surface-tension model SM5.0R with a relatively good overall performance is by far the fastest model regarding computation time, and that both SM2 and SM5.4A still require much less CPU time that the *ab initio* models COSMO-DFT and PCM-vdW.

MULTILINEAR REGRESSION COMBINING ELECTROSTATICS WITH MOLECULAR SURFACE AREA

A rough approximation for the nonelectrostatic component of ΔG_s (covering cavity formation and dispersion interaction) can be achieved by setting this term proportional to the molecular surface area of the solute, and by determining the relevant coefficient through regression with experimental data. Keeping in mind that the atom types differ in their propensity for dispersion interactions (i.e., their surface tension), a somewhat more refined but still simple approach would be to decompose the total surface area into its contributions from the individual atom types, which would allow the nonelectrostatic component to be written as linear combination of partial surface areas of the relevant atom types, with potential differences between the individual regression coefficients that would implicitely account for differences in their surface tensions.

Inclusion of the molecular surface area (calculated in the solution-phase geometry as optimized with COSMO-DFT/BPW-dnp) as a second variable in regression models with any of the electrostatic terms $\Delta G_s^{\text{MST-el}}$, $\Delta G_s^{\text{SCF-el}}$, $\Delta G_s^{\text{MP2-el}}$, ΔE_s^{BPW}, $\Delta G_s^{\text{SM2-ENP}}$ and $\Delta G_s^{\text{SM5.4A-ENP}}$ yields improved statistics only in the first two cases ($r^2_{\text{adj}} = 0.74$ and $r^2_{\text{adj}} = 0.91$, respectively). When offering the partial surface areas summarized separately for the atom types C, H, N, O, and halogen, the best statistics are achieved with the COSMO-DFT/BPW-dnp solvation energy and the surface area terms for nitrogen and oxygen, SA(N) and SA(O):

$$\begin{aligned} \log H = {} & 0.263(\pm 0.011) \cdot \Delta E_s^{\text{BPW}} \\ & + 0.0566(\pm 0.0092) \cdot \text{SA(N)} \\ & + 0.0099(\pm 0.0032) \cdot \text{SA(O)} \\ & + 9.32(\pm 0.22) \\ r^2_{\text{adj}} = {} & 0.96, \qquad \text{SE} = 0.40, \\ F_{3,36} = {} & 339.3, \qquad n = 40 \end{aligned} \tag{4}$$

In this equation, the positive signs of the regression coefficients of SA(N) and SA(O) are in accord with expectation: Assuming an overall destabilizing contribution from the sum of cavitation energy and dispersion interaction, greater surface areas should yield smaller negative solvation energies, and thus greater values for Henry's law constant.

With $\Delta G_s^{\text{SCF-el}}$ and again both SA(N) and SA(O), the resultant r^2_{adj} value is 0.94, with SE = 0.50, and again both regression coefficients are positive. By contrast, regression analysis with $\Delta G_s^{\text{MP2-el}}$ yields SA(C) and SA(O) as the only remaining surface area terms. In this case, however, both surface area regression coefficients have negative signs, and the respective statistics are only slightly better than for $\Delta G_s^{\text{MP2-el}}$ alone ($r^2_{\text{adj}} = 0.95$, SE = 0.47).

Overall, these findings suggest that there is a potential for further exploiting the simple approach of modeling cavitation energy and dispersion interaction by surface area terms that are specific for individual atom types. However, a substantially greater data set will be needed to derive better estimates for the individual regression coefficients.

Conclusions

For the data set of 40 aromatic compounds, the differences in performance between the individual continuum-solvation models are surprisingly large, regarding both the absolute values for the solvation energy and the relative trends. In a substantial number of cases, predictions of absolute Henry's law constants are not feasible with the presently used parameterization schemes, while scaling through linear regression results in r^2_{adj} values above 0.90 for MP2//PCM-vdW, COSMO-DFT, SM2, and SM5.0R with standard errors below and around 0.6 logarithmic units.

The significant differences between the nonelectrostatic components of the solvation energy according to SM2, SM5.4A, MST-AM1, and PCM-vdW suggest that these terms depend distinctly on the overall parameterizations of the individual continuum-solvation models. At present, a separate use of the currently available nonelectrostatic terms to quantify cavitation energy and dispersion interactions is not recommended. It follows that there is a need for developing improved methods to calculate these energy terms, which will be particularly relevant for more hydrophobic compound classes of higher environmental concern like polychlorinated biphenyls, dioxins, and furans, where dispersion interactions are likely to have a major impact on Henry's law constant.

COSMO-AM1, MST-AM1, and SM5.4A perform significantly better when omitting the six nitro compounds, and for amino and nitrilo derivatives significant differences are observed between the dipole moments of AM1 and DFT as well as between COSMO-AM1 and COSMO-DFT. These findings demonstrate apparent pitfalls of AM1 with nitrogen in various functional groups, which should be taken into account for future refinements of continuum-solvation schemes based on the AM1 level of quantum chemistry.

Regarding dipole moments, the MP2 correction generally lowers the values as predicted by HF/6-31G**, which is particularly pronounced for the aromatic nitro compounds. It indicates that for aromatic rings with NO_2 groups, the computational results are particularly sensitive to the level of considering electron correlation.

Acknowledgments

The Gaussian 94 module to perform solvation energy calculations according to the PCM-vdW scheme was generously provided by Dr. Maurizio Cossi, Professor Vincenzo Barone (Napoli, Italy), and Professor Jacopo Tomasi (Pisa, Italy), which is gratefully acknowledged. Moreover, the author thanks anonymous reviewers for helpful comments.

References

1. Tomasi, J.; Persico, M. Chem Rev 1994, 94, 2027.
2. Cramer, C. J.; Truhlar, D. G. In Lipkowitz, K. B.; Boyd, D. B., Eds.; Reviews in Computational Chemistry 6; VCH: New York, 1995, p. 1.
3. Mineva, T.; Russo, N.; Sicilla, E. J Comput Chem 1998, 19, 290.
4. Adamo, C.; Cossi, M.; Barone, V. J Comput Chem 1997, 18, 1993.
5. Schüürmann, G. In Chen, F.; Schüürmann, G., Eds.; Quantitative Structure–Activity Relationships in Environmental Sciences—VII; SETAC Press: Pensacola, 1997, p. 225.
6. Schüürmann, G.; Cossi, M.; Barone, V.; Tomasi, J. J Phys Chem A 1998, 102, 6706.
7. Schüürmann, G. J Chem Phys 1998, 109, 9523.
8. Schüürmann, G.; Somashekar, R. K.; Kristen, U. Environ Toxicol Chem 1996, 15, 1702.
9. Thomas, R. G. In Lyman, W. J.; Reehl, W. F.; Rosenblatt, D. H., Eds.; Handbook of Chemical Estimation Methods; American Chemical Society: Washington, DC, 1990.
10. Shiu, W.-Y.; Mackay, D. J Chem Eng Data 1997, 42, 27.
11. Peng, J.; Wan, A. Environ Sci Technol 1997, 31, 2998.
12. Howard, P. H.; Meylan, W. M. In Chen, F.; Schüürmann, G., Eds.; Quantitative Structure–Activity Relationships in Environmental Sciences—VII; SETAC Press: Pensacola, 1997, p. 185.
13. Cramer, C. J.; Truhlar, D. G. Science 1992, 256, 213.
14. Jorgensen, W. L.; Nguyen, T. B. J Comput Chem 1993, 14, 195.
15. Barone, V.; Cossi, M.; Tomasi, J. J Chem Phys 1997, 107, 3210.
16. Schüürmann, G. Environ Toxicol Chem 1995, 14, 2067.
17. Luque, F. J.; Bachs, M.; Orozco, M. J Comput Chem 1994, 15, 847.
18. Klamt, A.; Schüürmann, G. J Chem Soc Perkin Trans 2, 1993, 799.
19. Marten, B.; Kim, K.; Cortis, C.; Friesner, R.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. J Phys Chem 1996, 100, 11775.
20. Tawa, G. J.; Topol, I. A.; Burt, S. K.; Caldwell, R. A.; Rashin, A. A. J Chem Phys 1998, 109, 4852.
21. Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. J Phys Chem 1996, 100, 16385.
22. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. J Phys Chem B 1997, 101, 7147.
23. Andzelm, J.; Kölmel, C.; Klamt, A. J Chem Phys 1995, 103, 9312.
24. Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. J Phys Chem A 1998, 26, 5074.
25. Bachs, M.; Luque, F. J.; Orozco, M. J Comput Chem 1994, 15, 446.
26. Beauman, J. A.; Howard, P. H. Physprop Database; Syracuse Research Corporation: Syracuse, NY, 1996.
27. Sorbe, G. Sicherheitstechnische Kenndaten; Ecomed, Landsberg/Lech, Germany, 1983–1998.
28. Yalkowsky, S. H.; Dannenfelser, R.-M. AQUASOL dATAbASE of aqueous solubility; University of Arizona: Tucson, AZ, 1990, 5th ed.
29. Stephenson, R. M. J Chem Eng Data 1994, 39, 225.
30. Kuyper, L. F.; Hunter, R. N.; Ashton, D.; Merz, K. M.; Kollman, P. A. J Phys Chem 1991, 95, 6661.
31. Yalkowsky, S. H.; Valvani, S. C. J Pharm Sci 1980, 69, 912.
32. Isnard, P.; Lambert, S. Chemosphere 1989, 18, 1837.
33. Umweltbundesamt, Ed. Dokumentation wässergefährdender Stoffe: Datenblattsammlung zur Verwaltungsvorschrift wassergefährdender Stoffe (VwVwS) nach §19g WHG; Hirzel: Stuttgart, Germany, 1996–1998.
34. Eastin, W. NTP Chemical Health and Safety Data, http://ntp-db.niehs.nih.gov/Main_pages/Chem-HS.HTML.
35. CambridgeSoft Corp. (CS):ChemFinder WebServer, http://chemfinder.camsoft.com/.
36. Wauchope, R. D.; Buttler, T. M.; Hornsby, A. G.; Augustijn-Beckers, P. W. M.; Burt, J. P. Rev Environ Contam Toxicol 1992, 123, 1.
37. Rippen, G., Ed. Handbuch Umweltchemikalien; Ecomed: Landsberg/Lech, Germany, 1989–1998.
38. Stephenson, R. M. J Chem Eng Data 1993, 38, 634.
39. SYBYL Molecular Modelling Software 6.4; Tripos Associates Inc.: St. Louis, MO, 1996.
40. MOPAC 93, Rev. 2. Fujitsu Ltd., 9-3, Nakase 1-Chome, Mihama-ku, Chiba-city, Chiba 261, Japan; Stewart Computational Chemistry, 15210 Paddington Circle, Colorado Springs, CO 80921, 1994.
41. AMSOL 4.0, QCPE 606, Bloomington, IN 47405, 1993; AMSOL 6.5.3, Hawkins, G. D.; Giesen, D. J.; Lynch, G. C.; Chambers, C. C.; Rossi, I.; Storer, J. W.; Li, J.; Zhu, T.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, T. G. University of Minnesota (1998), based in part on AMPAC 2.1 by D. A. Liotard, E. F. Healy, J. M. Ruiz, and M. J. S. Dewar, and on the EF routines by Frank Jensen.
42. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.;

Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian 94, Revision D.4; Gaussian Inc.: Pittsburgh, PA, 1995.

43. Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. Chem Phys Lett 1996, 225, 327.
44. DMol Version 960; Biosym Technologies: San Diego, CA, 1996.
45. Cossi, M.; Mennucci, B.; Pitarch, J.; Tomasi, J. J Comput Chem 1998, 19, 833.
46. Becke, A. D. J Chem Phys 1988, 88, 2547.
47. Perdew, J. P.; Wang, Y. Phys Rev B 1992, 45, 13244.
48. Lee, C.; Yang, W.; Parr, R. G. Phys Rev B 1988, 37, 785.
49. Schreiner, A. C.; Baker, J.; Andzelm, J. W. J Comput Chem 1997, 18, 775.
50. Martell, J. M.; Goddard, J. D.; Eriksson, L. A. J Phys Chem A 1997, 101, 1927.
51. Klamt, A.; Jonas, V. J Chem Phys 1996, 105, 9972.
52. MOLSV; QCPE 509: Bloomington, IN, 1985.
53. Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. J Comput Aided Mol Design 1995, 9, 87.